

- 1 -

# DATABASE MANAGEMENT SYSTEM WITH REBALANCE ARCHITECTURES

## CROSS-REFERENCES TO RELATED APPLICATION

This application relates to copending Patent Application Serial No. 09/702,351 entitled "DATABASE MANAGEMENT METHODS AND EQUIPMENT, AND DATABASE MANAGEMENT PROGRAM STORAGE MEDIA" filed by Nobuo Kawamura on October 31, 2000.

## BACKGROUND OF THE INVENTION

The present invention relates to a database for data management based on table, and more particularly to a database system suitable to rebalance  
5 data among shared database files at the time of changing the number of storages for storing therein table data as divided.

B. Bergsten et al., "Overview of Parallel Architectures for Databases", THE COMPUTER JOURNAL,  
10 Vol. 36, No, 8, 1993, pp. 734-740 teaches an architecture wherein the processing load of a relational database management system (RDBMS) is distributed into a plurality of processors for parallel execution. D. J. DeWitt et al., Parallel Database  
15 Systems: The future of Database Processing or a Passing Fad?, SIGMOND RECORD, Vol. 19, No. 4, December 1990, pp. 104-112 discloses a database system wherein table data are stored as divided in a plurality of disks (storages) for distribution of disk access in RDBMS.

09987839-11601

However, as the amount of data increases, this will involve lack of capacities of the storages or performance deterioration. To avoid this, it becomes necessary to increase the number of disks for storage of table data by increasing the number of divisions.

Accordingly, even when the data rebalance technique in the aforementioned disk array systems is

applied to a database system, it is difficult to effectively execute the data rebalance processing and data access parallelly or concurrently. That is, it is required in the database that the data rebalance and  
5 data access be executed concurrently while eliminating the need for sharing information such as a position during the data rebalance.

The problem in the prior art is that database processing such as search, update, delete or insert to  
10 a table cannot be executed during data rebalance execution caused by changing the number of storages for storing tables.

#### SUMMARY OF THE INVENTION

It is therefore an object of the present  
15 invention to provide a database system, database management method and program, which, even during data rebalance execution of table data, can perform acceptance and execution of a database processing request such as search, update, delete or insert  
20 concurrently with the rebalance execution and with improved operability and performance of a database.

In accordance with an aspect of the present invention, the above object is attained by providing a database management system which includes a storage  
25 having a first storage area corresponding to a plurality of storages for saving a plurality of data items and a second storage area corresponding to a

6967639-11601

storage to be added to or disconnected from the plurality of storages, an acceptance section connected to the storage for accepting a data processing request, the data processing request including data processing  
5 in the plurality of storages and data rebalance between the storages, and a plurality of executors connected to the acceptance section for sequentially executing at least any of the data processing in the plurality of storages and the data rebalance.

10           The storage may have a storage area correspondence table indicative of combinations of predetermined data items to be shared by the plurality of storages and to be sharedly stored thereby according to the request of addition or disconnection to cause  
15 the data rebalance between the storages.

          The acceptance section has a during-rebalance flag indicating that 'during rebalance' of the plurality of storages by addition or disconnection to or from the plurality of storages. According to a data  
20 processing request to the data items stored in the plurality of storages, the acceptance section can refer to the during-rebalance flag and reflect data update even on the storages subjected to the data rebalance.

          The system in accordance with the invention  
25 of the present application may include means, in response to a rebalance request of data to be rebalanced in a storage added according to an addition request, for adding data position information to data

09087839-44504

before subjected to the rebalance execution by the data  
rebalance request in the plurality of storages, and  
means for deleting the data added with the data  
position information before subjected to the rebalance  
5 execution in response to the rebalance request after  
the rebalance execution by the rebalance request.

In accordance with another aspect of the  
present invention, there is provided a database  
management method which can be realized in the above  
10 system. The method includes steps of allocating a  
plurality of data items in a table to a plurality of  
storages for their storage; storing the table data in  
the storages determined according to a specified  
division rule (method); when a request is issued to add  
15 a storage for storage of the table data, determining  
part of the data stored in the existing storages which  
is to be moved to the addition storage with use of  
information about the existing and additional storages  
according to the specified division method and  
20 performing data rebalance operation to move all the  
determined data to the additional storage; when a  
request for search, update delete to the table is  
issued during the data rebalance execution, searching  
for, updating or deleting the table data stored in the  
25 existing storages and in the additional storage; and  
when a request of insert in the table is issued,  
inserting the data determined by the specified data  
division method in the storages together with

information about existing and additional storages. In accordance with a further aspect of the present invention, there is provided a parallel database method which includes steps of allocating a plurality of data items in a table as table data to a plurality of storages for their storage and storing the table data in storages determined by a specified division method; when a request is issued to search, update or delete the table is issued, performing parallel operation over the respective storages; when a request is issued to add a storage for storage of the table data, determining part of the data stored in existing storage parts which is to be moved to the additional storage according to a specified division method with use of information about the existing and additional storages; performing data rebalance operation to move all the determined data to the additional storage; when a request is issued to search, update or delete the table during the data rebalance operation, performing parallel operation in the existing storages over the table data stored therein to search for, update or delete the table data; after completing all the operations of the existing storages, performing parallel operation over the respective storages to search, update or delete the table in the additional storage; and when a request is issued to insert data in the table, inserting the data determined by the specified data division method in the existing and

In accordance with a further aspect of the present invention, there is provided a database method which includes steps of allocating a plurality of data items in a table as the table data to a plurality of storages for their storage; storing the table data in the storages determined by a specified division method; when a request is issued to add a storage for storage of the table data therein, determining part of the data stored in existing storages which is to be moved to the additional storage according to the specified division method with use of information about the existing and additional storages, copying the determined data from the existing storages to the additional storage, and adding position information of the copied data on the additional storage to the data of the storages as a copy source; when, the copy of all the data to be moved to the additional storage is completed, performing data rebalance operation to delete all the data of the copy source; when a request is issued to search the table during the data rebalance execution, searching the existing storages for associated data stored therein; when a request is issued to update or delete the table, updating or deleting associated data stored in the existing storages; when position information indicative of a copy target to the additional storage is added to the data to be updated or deleted, updating or deleting

5

10

## BRIEF DESCRIPTION OF THE DRAWINGS

15

Fig. 2 is a block diagram of an exemplary structure of a hardware section in the database system of Fig. 1;

20

25

Fig. 5 is a block diagram of a first exemplary detailed arrangement of the database



management system in Fig. 1; prior to the rebalance operation thereof;

Fig. 6 shows an example of a stock table to be managed by the database management system of Fig. 5;

5        Fig. 7 is a flowchart for explaining a first example of processing operation of a DBMS acceptance section in Fig. 5;

Fig. 8 is a flowchart for explaining a second example of processing operation of the DBMS acceptance  
10 section in Fig. 5;

Fig. 9 is a flowchart for explaining a third example of processing operation of a DBMS acceptance section in Fig. 5;

Fig. 10 is a diagram for explaining the first  
15 operational example of the database management system of Fig. 5 associated with the rebalance operation;

Figs. 11A and 11B show a flowchart for explaining an example of the rebalance operation of the database management system of Fig. 10;

20        Fig. 12 is a block diagram of an exemplary detailed arrangement of the database system of Fig. 1 after the rebalance operation;

Fig. 13 is a block diagram of a second exemplary detailed arrangement of the database  
25 management system in Fig. 1 prior to the rebalance operation;

Fig. 14 is a flowchart for explaining an example of search execution target determining

09037339-111504

operation of a DBMS acceptance section in Fig. 13;

Fig. 15 is a flowchart for explaining an example of update/delete execution target determining operation of the DBMS acceptance section in Fig. 13;

5 Fig. 16 is a flowchart for explaining an example of update execution target determining operation of the DBMS acceptance section in Fig. 13;

Fig. 17 is a flowchart for explaining an example of delete execution target determining operation of the DBMS acceptance section in Fig. 13;

Fig. 18 is a flowchart for explaining an example of insert execution target determining operation of the DBMS acceptance section in Fig. 13;

Fig. 19 is a flowchart for explaining another example of insert execution target determining operation of the DBMS acceptance section in Fig. 13;

Fig. 20 is a diagram of another exemplary arrangement of the database management system of Fig. 13 associated with the rebalance operation;

20 Figs. 21A and 21B shows a flowchart for explaining an example of rebalance operation of the database management system of Fig. 20;

Figs. 22A and 22B are diagrams for explaining a second example of accepting and executing a database processing request in the rebalance operation of the database system of Fig. 1; and

Figs. 23 to 25 are diagrams for explaining the data processing of the rebalance operation

00567339.11607

## DESCRIPTION OF THE EMBODIMENTS

Fig. 1 is a block diagram of an exemplary arrangement of a database system in accordance with an embodiment of the present invention, Fig. 2 is a block diagram of an example of a hardware section in the database system of Fig. 1, Fig. 3 shows exemplary registration contents of a table of hash function value storage position correspondence in Fig. 1, and Fig. 4 shows diagrams for explaining a first example of accepting and executing a database processing request in rebalance operation of a database management system in Fig. 1.

As shown in Fig. 2, a database management system (which is also shown by DBMS in the drawings) 300 forming one section of the database system of the present invention, which includes a central processing unit (CPU), an input device, a display unit and so on, functions to perform database processing such as data search, update, delete, or insert (add) on the basis of a request from a plurality of user terminals 330a to 330c each having a computer function. The database management system 300 includes a DBMS acceptance node 400 and DBMS execution nodes 410a to 410c, which have

The DBMS acceptance node 400 has a CPU 401 and a main memory 402 and also is connected to an external storage 318 in the form of a hard disk drive (HDD) or the like. When the CPU 401 executes a DBMS acceptance program 310a stored in the main memory 402, the DBMS acceptance section 310 in Fig. 1 can perform respective functions in the DBMS acceptance operation.

20           The DBMS acceptance program 310a and DBMS  
execution programs 320a to 320c are usually stored in a  
recording medium including an optical disc such as CD-  
ROM (compact disc-read only memory) or DVD (digital  
video disc/digital versatile disc). These programs are  
25 installed from these recording medium to the external  
storage 318 and external storages 325a to 325c to be  
loaded into the main memory 402 and main memories 412a  
to 412c.

A DBMS acceptance section 310 and the DBMS execution programs 320a to 320c in the DBMS 300 shown in Fig. 1, which correspond to the DBMS acceptance node 400 and DBMS execution nodes 410a to 410c shown in Fig. 5 2, are interconnected by the network 301.

The DBMS acceptance section 310 has a request acceptor 311, an analyzer 312, a processing procedure generator 313, an execution destination decider 314, an execution request/result receiver 315, a returner 316, 10 and a rebalance executor 102. The DBMS executors 320a to 320c have searches 321a to 321c, updates 322a to 322c, deletes 323a to 323c, and inserts 324a to 324c respectively.

The DBMS acceptance section 310 accepts from 15 user terminals a processing request 330 such as data processing and rebalance operation via the request acceptor 311, analyzes the request in the analyzer 312 therein, and generates a processing procedure to realize the request in the processing procedure 20 generator 313.

Then the execution destination decider 314 determines which one of the DBMS executors should execute the generated processing procedure on the basis of data division conditions, the execution 25 request/result receiver 315 issues an execution request to the determined DBMS executor and receives its execution result from the DBMS executor. The received result is returned from the returner 316 to the user.

09087839 1150  
10571 528/8660

The analyzer 312, processing procedure generator 313 and execution destination decider 314 refer to a hash function value storage position correspondence table 317 or table definition information 319 stored in the external storage 318, and acquire information for determination of a data storage destination or information about the table.

Stored in the table definition information 319 are, for example, storage area information 340 for management of storages having the table stored as divided therein, addition area information 341 for management of a storage for storage area addition, and a during-rebalance flag 342 indicative of 'during rebalance operation'. Details of respective constituent elements of the table definition information 319 will be explained later in connection with a specific example with reference to Figs. 5 to 12.

The DBMS executors 320a to 320c, in response to an instruction of search, update, delete or insert received from the DBMS acceptance section 310, activate the searches 321a to 321c, updates 322a to 322c, deletes 323a to 323c or inserts 324a to 324c, execute its processing procedure, and return the execution result to the DBMS acceptance section 310.

During the execution of the processing procedure, the DBMS executors 320a to 320c access table data 326 stored in the external storages 325a to 325c.

FOR OFFICIAL USE ONLY

In such processing, when the execution destination decider 314 of the DBMS acceptance section 310 selects a plurality of executors (320a to 320c), the execution request/result receiver 315 of the DBMS acceptance section 310 may issue an execution request to all the selected DBMS executors (320a to 320c) simultaneously to execute parallel processing in the respective DBMS executors (320a to 320c).

Explanation will now be made as to the data dividing operation in the DBMS 300. In this example, when data are stored as divided, a hash function is applied to the data to be stored to generate values of from 0 to 9, and the data are distributed to respective storage destinations according to the generated values.

The table 317 showing its details in Fig. 3 is used to determine a storage position based on the hash function value. In this example, data is divided into up to 10 parts according to the number of external storages connected. The table also shows information about storage areas where data is stored at the time of the division by the hash function values.

For example, when data is stored as divided into two storages, as shown items of "area 0" and "area 1" at the time of two divisions in the hash function value storage position correspondence table 317, data about hash function values of 0, 4, 6, 7 and 9 are stored in the "area 0", and data about hash function values of 1, 2, 3, 5 and 8 are stored in the "area 1".

When a storage area was added, even data to be moved by the rebalance executor 102 can be found from the hash function value storage position correspondence table 317. When a single storage is added and two divisions are changed to three divisions, for example, "area 2" is added and data of hash function values of 2, 7 and 3 are stored in "area 2", as in the case of three divisions in the hash function value storage position correspondence table 317. Thus it is only required to move the data of the hash function values of 2, 7 and 3 so far stored in "area 0" and "area 1" to "area 2" as an addition area.

Similarly, in the case of changing from two divisions to five divisions, "area 2", "area 3" and "area 4" are added, data of hash function values of 2 and 7; 3 and 6; and 4 and 5 are moved to "area 2", "area 3" and "area 4" respectively.

For data moving operation when such a storage area was added, that is, for rebalance and acceptance of a request of database processing (such as search, update, delete, or insert); the database management system 300 in the database system having such an arrangement as shown in Fig. 1 copes with it in such a manner as to be explained below.

More specifically, when it is required to add a storage for table data such as an external storage 325c external storage, the rebalance executor 102 in the DBMS acceptance section 310 refers to the hash

090703041604



5

10

20

completing the processing of the existing storages, the system sequentially performs the search, update and delete operations over the additional storage and external storage 325c. Further, when it is required to  
5 insert data in the table data, the system refers to the hash function value storage position correspondence table 317 including information on the existing storages (external storages 325a and 325b) and information on the additional storage (external storage  
10 325c), and inserts the data in one of the storages (external storages 325a, 325b and 325c) determined by the execution destination decider 314.

Explanation will be made in more detail by referring to Fig. 4 as to how to control the data  
15 moving operation by the database management system 300 in Fig. 1 in accordance with the present invention when a storage area is added, that is, as to operation of accepting and executing a database processing request during rebalance operation.

20 Explanation will be made as to the rebalance operation as well as the search, update, delete and insert operations during the rebalance operation when there is a table wherein data indicative of the hash function values of 0, 4, 6, 7 and 9 are stored as  
25 divided in a storage 120 and data indicative of the hash function values of 1, 2, 3, 5 and 8 are stored as divided in a storage 130 as shown in Fig. 4A.

In the arrangement shown in Fig. 4A,

09907539.44604

5 is carried out in the storage 130.

10 data to be moved to the storage 140 from data stored in  
the storages 120 and 130 and moves the selected data to  
the storage 140 to provide hash function values of 7, 2  
and 3.

15 such a procedure as to first search for data of the  
storages 120 and 130 to be moved, insert it in the  
storage 140 and then delete the data of the storages  
120 and 130.

20 after the search for the data to be moved until the  
deletion of the old data already moved, the data is  
locked so that another user is prohibited from  
searching for, updating or deleting the data.

25 execution of the rebalance operation 102a, the data  
operation 100 of search, update or delete is carried  
out in both existing and additional areas. That is,  
the data operation is carried out in the storages 120,

5 specifically stated by the user.

additional area to be moved.

15 operation releases the lock, he can resume the  
operation so far awaited.

```

20  access to the data is locked, until the lock is
    released. Once the lock is released, however,
    resume the rebalance operation so far awaited.

```

updated in its value or deleted.

25                    During the execution of the rebalance  
operation 102a, data to provide a hash function value  
of 7 is stored in the storage 120 or 140 and data to  
provide hash function values of 2 and 3 is stored in

the storage 120 or 140. However, since the operation 100 of search, update or delete refers to both the existing and additional areas, the operation 100 can carry out the search, update or delete operation regardless of whether the data is stored in the existing storages or additional storage.

Further, when data to be rebalanced is inserted in a rebalance destination, excessive rebalance operation to the existing areas can be avoided.

Through the above operation, it becomes possible to accept and concurrently execute search, update, delete and insert processing requests to table data during the rebalance operation.

15                Such processing operation will be explained  
in detail in connection with specific examples of Figs.  
5 to 12.

Fig. 5 is a block diagram of a first detailed example of the arrangement of the database management system in Fig. 1 prior to its rebalance operation, Fig. 6 shows exemplary contents of a stock table to be managed by the database management system of Fig. 5, Fig. 7 is a flowchart for explaining a first example of the processing operation of a DBMS acceptance section in Fig. 5, Fig. 8 is a flowchart for explaining a second example of the processing operation of the DBMS acceptance section in Fig. 5, Fig. 9 is a flowchart for explaining a third example of the processing operation

of the DBMS acceptance section in Fig. 5, Fig. 10 is a diagram for explaining the arrangement and operation of the database management system in Fig. 5 associated with the rebalance operation, Fig. 11 is a flowchart for explaining an exemplary rebalance operation of the database management system of Fig. 10, and Fig. 12 is a block diagram of an exemplary specific arrangement of the database system after the rebalance operation.

A DBMS acceptance section 600 and DBMS executors 610a to 610c in the database management system in the database system of Fig. 5 have substantially the same functions as the DBMS acceptance section 310 and DBMS executors 320a to 320c in the database management system 300 in Fig. 1.

Stored in a storage 630 connected to the DBMS acceptance section 600 are a hash function value storage position correspondence table 635 having the same contents as the hash function value storage position correspondence table 317 shown in Fig. 3 as well as table definition information 631 defined for each table. The table definition information 631 has a storage area (shown by "additional area" in the drawings) 632 indicative of a table storage place, additional area information (shown by "additional area" in the drawing) 633, and a during-rebalance flag 634 indicative of "during rebalance".

Storages 620a and 620b connected to the DBMS executors 610a and 610b respectively have storage data

621 and 622 of stock tables in the form of a database. The DBMS executors execute search, update, delete or insert operation in response to an instruction received from the DBMS acceptance section 600.

5           In this example, it is assumed that the number of divisions is set at 10 as a maximum, a hash function used for the division is set to have values of residues obtained when identification numbers (product codes) of products in storage data 621 and 622 are  
10 divided by the maximum division number 10.

          As shown in Fig. 6, the database treated in this example is a stock table 500 associated with clothing which includes items of product code 510, product name 520, color 530, unit price 540 and stock  
15 amount 550.

          When the stock table 500 of Fig. 6 is applied to the database system of Fig. 5 and data is stored as divided in the storages 620a and 620b with use of the value of the product code 510, the distribution of the  
20 hash function value storage position correspondence table 635, i.e., the hash function value storage position correspondence table 317 of Fig. 3 in the two-division case is applied to the hash function value calculated for the value of the product code 510.

25           As a result, data indicative of hash function values of 0, 9, 4, 6 and 7 are stored in the storage 620a, and data indicative of hash function values of 1, 8, 5, 2 and 3 are stored in the storage 620b. Further,

09097039.4404  
7097.52560

storages A and B are defined in the storage area 632 of the table definition information 631 in the storage 630.

When a search, update or delete request is  
5 issued under this condition, the DBMS acceptance section 600 accepts the request at a request acceptor 601, analyzes the accepted request at an analyzer 602, generates a processing procedure at a processing procedure generator 603 to realize the analyzed  
10 request, and determines one or ones of the DBMS executors which execute(s) the generated processing procedure at an execution destination decider 604.

In details of the operation of the execution destination decider 604, as shown in Fig. 7, the  
15 decider first refers to the value of the during-rebalance flag 634 in the table definition information 631 of the storage 630 to judge whether or not the rebalance operation is being executed (step S701). When determining that the rebalance operation is not  
20 being executed, the decider determines the area defined in the storage area 632 as its execution destination (step S702).

In the example of Fig. 5, since the storages A and B are defined in the storage area 632, the DBMS  
25 executors 610a and 610b are determined as the execution destinations. The operation when the judgement result at the step S701 indicates that the rebalance operation is being executed will be explained later.

00007030-1104  
FOR F. 6202000



When the receiver determines that the rebalance operation is not being executed, it issues a simultaneous execution request to all the DBMS executors determined by the execution destination decider 604 (step S712) and receives execution results from the DBMS executors (step S713).

When an insert request is issued to the database management system in Fig. 5, the DBMS acceptance section 600 performs the operations of the request acceptor 601, analyzer 602 and processing procedure generator 603, and determines an execution destination as an insert destination at the execution



DBMS acceptance section 600 issues an insert execution request and receives its result from and at the execution request/result receiver 605, and returns the last-received result from the returner 606.

5           In the database management system of Fig. 5, when a storage 620c is added as a storage area of the stock table as shown in Fig. 10, a storage C is defined in the additional area information 633 of the storage 630, a rebalance operation request is generated, and  
10 the operation of a rebalance executor 1100 is started.

Explanation will be made as to details of the rebalance operation of the rebalance executor 1100 with reference to Fig. 11. More specifically, the rebalance executor set the during-rebalance flag 634 in the table  
15 definition information 631 of the storage 630 in Fig. 5 at "during execution of the rebalance operation" (step S1101), finds the numbers of divisions before and after the area addition by referring to the storage area information 632 and additional area information 633,  
20 and finds data to be moved by referring to the hash function value storage position correspondence table 635 (step S1102).

A single area is added to change the system from the two-division type to a three-division type in  
25 Fig. 10. Thus it will be appreciated that data indicative of the hash function values of 2, 7 and 3 is to be moved on the basis of the hash function value storage position correspondence table 635 (hash

09087839.4464

The rebalance executor 1100 issues a data search request to the DBMS executor having the to-be-moved data thus found and stored therein (step S1103). In Fig. 10, since data indicative of the hash function value of 7 to be moved is stored in the storage 620a and data indicative of the hash function values of 2 and 3 are stored in the storage 620b, the rebalance executor issues a search request to the DBMS executors 610a and 610b.

In Figs. 11A and 11B, the DBMS acceptance section 600 accepts the data corresponding to one case returned from the DBMS executors 610a and 610b (step S1104), calculates a hash function value based on the value of the accepted data, finds a move destination by referring to the hash function value storage position correspondence table 635 (step S1105), and issues an insert request to the DBMS executor as the move destination (step S1106).

25           The DBMS executor when receiving the insert request executes data insert operation and informs the DBMS acceptance section 600 of its insert completion (step S614). The DBMS acceptance section 600 accepts a

notification indicative of the data insert completion (step S1107) and issues a delete request to delete the insert original data (step S1108).

For example, in the case of the rebalance  
5 executor 1100 in Fig. 10, the DBMS executor 610a searches for data 1000 having a product code value of 677 and returns it to the DBMS acceptance section 600. Since the data move destination is the storage 620c, the rebalance executor 1100 issues an insert request to  
10 the DBMS executor 610c to store data 1001 in the storage 620c.

After the insert completion, the rebalance executor 1100 issues a request to the DBMS executor 610a to delete the data 1000 having a product code  
15 value of 677.

When the DBMS executor 610a receives the delete request from the rebalance executor 1100, a delete 613a in the DBMS executor 610a deletes the corresponding data. After completing the delete  
20 operation, the DBMS executor 610a informs the DBMS acceptance section 600 of the delete completion. The DBMS acceptance section 600 accepts the data delete completion notification (step S1109), and extracts next search data and examines the presence or absence of the  
25 next search data (step S1110).

In the presence of the next data, the DBMS acceptance section 600 accepts data at the steps S1110 to S1104 and repetitively performs the insert and

00007030-1101

delete operations until the data becomes null. When  
the next data becomes null and the move operation of  
all the data is fully terminated, the DBMS acceptance  
section 600 adds contents of the additional area  
5 information 633 in the storage area 632 (S1111),  
deletes contents of the additional area information 633  
(S1112), and erases the during-rebalance flag 634  
(S1113).

After the rebalance operation is completed in  
10 this way, storage data 621a is stored in the storage  
620a, storage data 622 is stored in the storage 620b,  
and storage data 623 is stored in the storage 620c, as  
shown in Fig. 12.

Explanation will next be made as to the  
15 operation of the system when a data search, update and  
delete request and a data insert request are issued  
during the rebalance operation, that is, as to the  
database management method according to the present  
invention.

20 Explanation will first be made in connection  
with a case where a data search, update and delete  
request was issued. In the DBMS acceptance section 600  
in Fig. 10, in this case, the request acceptor 601,  
analyzer 602 and processing procedure generator 603 .  
25 performs their operations and the execution destination  
decider 604 performs its execution destination  
determining operation. In the execution destination  
determining operation, the DBMS acceptance section 600

00007039-1101  
FOR F-400

At this time, if the rebalance operation is

15           For the execution destinations thus  
determined, the execution request/result receiver 605  
in the DBMS acceptance section 600 performs its  
execution request/result receiving operation, which  
will be explained by referring to Fig. 8.

20           That is, the DBMS acceptance section 600  
refers to the during-rebalance flag 634 in the table  
definition information 631 of the storage 630 and  
judges whether or not the rebalance operation is being  
executed (step S711). Since the rebalance operation is  
25 being executed, the DBMS acceptance section 600, on the  
basis of the execution destination determined at the  
step S703 in the execution destination determination  
process of Fig. 7, first issues an execution request to

When accepting execution results from all the DBMS executors as the execution request destinations (step S715), the DBMS acceptance section 600 issues a execution request to the DBMS executor associated with the storage defined in the additional area 633 (step S716) and receives an execution result from the DBMS executor subjected to the execution request (step S717).

In the example of Fig. 10, the storages 620a and 620b are defined in the storage area 632 in the table definition information 631 of the storage 630 and the storage 620c is defined in the additional area 633. Thus the DBMS acceptance section 600 first issues a search, update and delete execution request to the DBMS executors 610a and 610b, and after receiving all their results therefrom, the DBMS acceptance section issues the search, update and delete execution request to the DBMS executor 610c and receives its result.

When an insert request is issued during the  
25 rebalance operation, the DBMS acceptance section 600  
performs the operations of the request acceptor 601,  
analyzer 602, processing procedure generator 603 and  
execution destination decider 604 in Fig. 10, and the



That is, in the step S721 wherein the during-rebalance flag is referred to to judge whether or not

In the example of Fig. 10, the DBMS acceptance section 600 refers to the value of the product code for the data to be inserted and finds hash function values. When the hash function values are 0, 9, 4 and 6, the DBMS acceptance section selects the storage 620a. When the hash function values are 1, 8 and 5, the DBMS acceptance section selects the storage 620b. When the hash function values are 2, 7 and 3, the acceptance section selects the storage 620c.

Such data being searched for, updated,

deleted or inserted is locked, so that it is prohibited that another user or rebalance operation performs search, update or delete operation thereover, unless otherwise specifically stated by the user. Similarly, data to be subjected to a move operation for each data case during the rebalance operation is also locked, so that another user is prohibited from performing search, update or delete operation thereover until the move operation is completed.

10 All the storages having data stored therein have been checked in order to determine the execution destination of the search, update or delete operation in the present embodiment. When the search, update or delete operation is added with a condition, however, 15 the check can be carried out over only limited one of the storages as the execution destination. When the storages as the execution destination are limited in the search, update or delete operation during the rebalance operation, it is only required to perform 20 update or delete operation over the additional area, only when data to be rebalanced is stored in the storage of the execution destination.

Through the above operation, even when another user issues a search, update or delete 25 processing request during rebalance operation, the system can accept the search, update or delete processing request to node table data during the rebalance operation and sequentially execute the DBMS

0907090440944

Explanation will next be made as to a database system in accordance with another embodiment of the present invention. In this database system,

15                   And when a search request to the table data  
is issued during the execution of the data rebalance  
operation, the system performs searching operation over  
the data stored in the existing storages. When an  
update/delete request to the table data is issued  
20 during the rebalance operation, the system first  
processes the data stored in the existing storages as  
update/delete object. When position information  
indicative of a copy destination to the additional  
storage is added to the data to be updated and deleted,  
25 the system performs the same update/delete operation  
over the data of the copy destination.

Further, when an insert request to the table data is issued during the execution of the rebalance

operation and the data in question is to be moved to the additional storage, the system stores the insert data in the existing storages and also even in the additional storage as a move destination, and adds  
5 storage position information to the additional storage to the data stored in the existing storages.

The basic arrangement of the database system performing such operation as mentioned above is the same as that of the system of Fig. 1, or may be the  
10 same as that of the system of Fig. 2.

By referring to Figs. 13 to 21, explanation will be made as to the operation of a database system as a second embodiment which has such an arrangement as mentioned above, and wherein data division is based on  
15 the hash function value storage position correspondence table 317 shown in Figs. 1 and 3 and the stock table of Fig. 6 is applied to the database system of Fig. 1.

Fig. 13 is a block diagram of a second detailed example of the arrangement of the database  
20 management system in Fig. 1 prior to rebalance operation, Fig. 14 is a flowchart for explaining an example of search execution destination determining operation of a DBMS acceptance section in Fig. 13, Fig.  
15 is a flowchart for explaining an example of  
25 update/delete execution destination determining operation of the DBMS acceptance section in Fig. 13, Fig. 16 is a flowchart for explaining an example of update execution destination determining operation of

09087828-44604  
TOP SECRET

In the database system of Fig. 13, when a search request is issued, a DBMS acceptance section 600a performs respective operations of the request acceptor 601, analyzer 602 and processing procedure generator 603, and determines an execution destination through the operation of the execution destination decider 604. In the execution destination determining operation of this example, the areas defined in the storage area 632 are merely determined to be the execution destinations as detailed in Fig. 14 (step

S1301).

In the example of Fig. 13, since the storages A and B are defined in the storage area 632, the DBMS executors 610a and 610b are determined as the execution destinations. And the DBMS acceptance section 600a issues a search execution request from the execution request/result receiver 605 to the execution destinations thus determined, receives their results, and returns the last-received result thereto from the returner 606.

Even with respect to data update or delete, when its request is issued, the DBMS acceptance section 600a performs the respective operations of the request acceptor 601, analyzer 602 and processing procedure generator 603, and determines an execution destination through the operation of the execution destination decider 604. In the execution destination determining operation, as in the search request, the areas defined in the storage area 632 are merely determined as execution destinations as shown in Fig. 15 (step S1401).

In the example of Fig. 13, since the storages A and B are defined in the storage area 632, the DBMS executors 610a and 610b are determined as the execution destinations. And the DBMS acceptance section 600a issues an update/delete execution request from the execution request/result receiver 605 to the execution destinations thus determined, receives their results,

The DBMS executors 610a and 610b search for the data to be updated/deleted (steps S1501 and SA1601) and execute the update/delete operation through the operations of updates 612a and 612b and deletes 613a and 613b (steps S1502 and S1602) as shown in Figs. 16 and 17.

Thereafter or in the steps S1503 and S1603, when the position information is not added, the DBMS executor extracts the data to be next updated/deleted and judges whether or not the next data is present (steps S1504 and S1604). In the presence of the next data, the DBMS executor repeats the operations of the steps S1501 and S1601 and subsequent steps until the data is fully searched for. In the absence of the next data, when the searching operation of all the data is completed, the DBMS executor informs the DBMS acceptance section 600a of the execution completion (steps S1505 and S1605).

When an insert request is issued, the DBMS acceptance section 600a in Fig. 13 performs the operations of the request acceptor 601, analyzer 602 and processing procedure generator 603, and determines  
5 an insert destination through the operation of the execution destination decider 604. In this example, the insert destination determining operation is as shown in Fig. 18.

More specifically, the DBMS acceptance  
10 section first refers to the during-rebalance flag 634 and judges whether or not the rebalance operation is being executed (step S1701). In this case, the DBMS acceptance section determines that the rebalance operation is not being executed, finds a hash function  
15 value from the data to be inserted, refers to the hash function value storage position correspondence table 635 (hash function value storage position correspondence table 317 in Figs. 1 and 3), and selects an insert destination from the areas defined in the storage area  
20 632 on the basis of the hash function value (step S1702).

In the example of Fig. 13, the DBMS acceptance section refers to the value of a product code for the insert data and finds its hash function  
25 value. When the hash function values are 0, 9, 4, 6 and 7, the DBMS acceptance section selects the storage 620a. When the hash function values are 1, 8, 5, 2 and 3, the DBMS acceptance section selects the storage



The DBMS acceptance section 600a issues an insert execution request from the execution request/result receiver 605 to the selected insert destination and receives its result therefrom. The execution request issuance and result reception are carried out according to a procedure as shown in Fig. 19.

As in the case of the example of Fig. 10, when the storage 620c as a storage area of the stock table is added to the database system of Fig. 13 as shown in Fig. 20, a storage C is defined in the additional area 633 of the storage 630 and the operation of a rebalance executor 1901 is started by a

As in the case of the example of Fig. 10, when the storage 620c as a storage area of the stock table is added to the database system of Fig. 13 as shown in Fig. 20, a storage C is defined in the additional area 633 of the storage 630 and the operation of a rebalance executor 1901 is started by a

rebalance operation request.

In the rebalance operation, as shown in Fig. 21, DBMS acceptance section first sets the during-rebalance flag 634 in Fig. 13 in 'during rebalance' during the execution of the rebalance operation (step S2001), refers to the storage area information 632 and additional area information 633, finds the number of divisions before the area addition and the number of division after the area division, refers to the hash function value storage position correspondence table 635 (hash function value storage position correspondence table 317 in Figs. 1 and 3), and finds data to be moved (step S2002).

In the example of Fig. 20, since a single area is added to the two division tables to form three division tables, it will be appreciated that data having hash function values of 2, 7 and 3 is to be moved from the hash function value storage position correspondence table 317 of Fig. 3.

Next, the DBMS acceptance section issues a data search request to the DBMS executor having the move data found at the step S2002 (step S2003). In the example of Fig. 20, since the hash function value of 7 for the move data is stored in the storage 620a and the hash function values of 2 and 3 are stored in the storage 620b, the DBMS acceptance section issues the search request to the DBMS executors 610a and 610b.

The DBMS executors 610a and 610b when

receiving the search request from the DBMS acceptance  
section 600a search for one case of data and return its  
result to the DBMS acceptance section 600a. The DBMS  
acceptance section 600a receives one case of result in  
5 the rebalance executor 1901 (step S2004), calculates a  
hash function value based on the received data value,  
refers to the hash function value storage position  
correspondence table 635 (hash function value storage  
position correspondence table 317 in Figs. 1 and 3),  
10 finds a move destination (DBMS executor 610c in Fig.  
20) (step S2005), and issues an insert request to the  
move destination (step S2006).

The DBMS executor 610c when receiving the  
insert request inserts the data in the storage 620c  
15 through the operation of an insert 614c, and informs  
the DBMS acceptance section 600a of data insert  
position information and insert operation completion.

The DBMS acceptance section 600a receives the  
data insert position information and insert operation  
20 completion at the rebalance executor 1901 (step S2007),  
and adds move-destination data insert position  
information to the move source data (step S2008).

In the example of Fig. 20, the DBMS executor  
610a searches for data ('trainer') 1911 having a  
25 product code value of 677, returns it to the DBMS  
acceptance section 600a. Since the data move  
destination is the storage 620c, the DBMS acceptance  
section 600a issues an insert request to the DBMS

executor 610c and data 1921 is stored in the storage 620c. And after completion of the insert operation, the move-destination position information 1921 is added to the data 1911 of the move source of the DBMS

5 executor 610a.

After the operation of the step S2008, the rebalance executor 1901 of the DBMS acceptance section 600a in Fig. 13 reads next search data and judges whether or not the next data is present (step S2009).

10 In the presence of the next data, the rebalance  
 executor repeats the data acceptance, move-destination  
 specifying, data inserting and move-destination  
 position information adding operations of the steps  
 S2004 to S2008 until the operation of the data is fully  
 15 completed.

When the next data becomes null move operation of all the data is completed in a step S2009, the DBMS acceptance section issues a delete request of the data added with the move-destination position information to the DBMS executor having so far stored therein the move data (step S2010).

The DBMS executor when receiving the delete request searches for and deletes the target data. After completion of the delete operation, the DBMS executor informs the DBMS acceptance section 600a of the delete completion. Thus the DBMS acceptance section 600a receives the delete completion of the data (step S2011), adds the additional area information 633

to the storage area information 632 (step S2012), deletes the additional area information 633 (step S2013), and erases the during-rebalance flag (step S2014).

5                   When the rebalance operation is completed in this way, data 621 (stock table) so far stored in the storage 620a and data 622 (stock table) so far stored in the storage 620b in Fig. 20 are rebalanced as in the case of Fig. 12 so that the storage data 621a is placed  
10 in the storage 620a, the storage data 622a is placed in the storage 620b, and the storage data 623 is placed in the storage 620c.

Explanation will then be made in connection with a case where a search or update, delete and insert  
15 processing request is issued during such rebalance operation.

When a search processing request is issued during the rebalance operation, first of all, the DBMS acceptance section 600a in Figs. 13 and 20 accepts a  
20 request at the request acceptor 601, analyzes the request at the analyzer 602, generates a processing procedure at the processing procedure generator 603 and determines an execution destination at the execution destination decider 604.

25                   In the execution-destination determining operation, the areas defined in the storage area 632 of the table definition information 631 of the storage 630 in Fig. 20 are determined as the execution destinations

In the example of Fig. 20, since the storages A and B are defined in the storage area 632, the DBMS acceptance section issues a search execution request to the DBMS executors 610a and 610b.

Next, when an update or delete processing request is issued during the rebalance operation, the DBMS acceptance section 600a accepts the request at the request acceptor 601, analyzes the request at the analyzer 602, generates a processing procedure at the processing procedure generator 603, and determines an execution destination at the execution destination decider 604.

25           In the example of Fig. 20, since the storages  
A and B are defined in the storage area 632, the DBMS  
acceptance section issued an update or delete execution  
request to the DBMS executors 610a and 610b.

The DBMS executors 610a and 610b search for (steps S1501 and S1601), updates or deletes (steps S1502 and S1602) data specified by the deletes 613a and 613b, as shown in Figs. 16 and 17.

5           Further, DBMS acceptance section refers to the update or delete data, judges whether or not position information indicative of a copy destination is added to the update or delete data (steps S1503 and S1603). When the position information is added, the  
10 DBMS acceptance section updates or deletes data at the copy destination (steps S1506 and S1606).

          Thereafter or in the steps S1503 and S1603, if the position information is not added, the DBMS acceptance section reads next data specified to be  
15 updated or deleted and judges whether or not the next data is present (steps S1504 and S1604).

          In the presence of the next data, the DBMS acceptance section repeats the operations of the steps S1501 and S1601 and subsequent steps until all the data  
20 is searched for. In the absence of the next data, when the searching operation of all the data is completed, the DBMS executors 610a to 610b inform the DBMS acceptance section 600a of the execution completion (steps S1505 and S1605).

25           When the DBMS acceptance section 600a receives the execution completion notification from the DBMS executors 610a and 610b in this way, the DBMS acceptance section returns the last received result.

5           When an insert request is issued, the DBMS acceptance section 600a in Figs. 13 and 20 performs the respective operations of the request acceptor 601, analyzer 602 and processing procedure generator 603, and determines an insert destination at the execution destination decider 604. In this example, the insert destination determining operation is as shown in Fig. 18.

More specifically, the DBMS acceptance section first refers to the during-rebalance flag 634 and determines whether or not the rebalance operation is being executed (step S1701). Because the rebalance operation is being executed, the DBMS acceptance section finds a hash function value based on the data to be inserted (step S1701), refers to the hash function value storage position correspondence table 635 (hash function value storage position correspondence table 317 in Figs. 1 and 3), determines an insert destination (1) (to be arbitrarily inserted) from the storage area 632 on the basis of the hash function value (step S1703), and determines an insert destination (2) (to be arbitrarily inserted) from the areas defined in the storage area 632 and the area defined in the additional area 633 (step S1704).



In the example of Fig. 20, when the hash function values are 0, 9, 4, 6 and 7, the storage 620a is selected as the insert destination (1); while, when the hash function values are 1, 8, 5, 2 and 3, the storage 620b is selected.

When the hash function values are 0, 9, 4 and 6, the storage 620a is selected. When the hash function values are 1, 8 and 5, the storage 620b is selected. When the hash function values are 2, 7 and 3, the storage 620c is selected.

The DBMS acceptance section 600a issues an insert execution request from the execution request/result receiver 605 to the insert destination thus selected, and receives their results. The issuance of the execution request and acceptance of the results are as shown in Fig. 19.

More specifically, the execution request/result receiver 605 in Fig.20 first refers to the during-rebalance flag 634 in the table definition information 631 of the storage 630 and judges whether or not the rebalance operation is being executed (step S1801).

Since the rebalance executor is now being executed, the DBMS acceptance section examines whether the storage as the insert destination (1) coincides with the storage as the insert destination (2) (step S1804). When finding a coincidence therebetween, the DBMS acceptance section issues an insert execution

request to the determined insert destination (step S1805), and receives its result (step S1806).

When insert destination (1) is different from the insert destination (2) in the examination result of the step S1804, the DBMS acceptance section issues an insert execution request first to the insert destination (2) (step S1807) and, after completion of the insert operation at the associated DBMS executor, receives insert destination data storage position information therefrom (step S1808).

Next, the DBMS acceptance section adds the data storage position information received at the time of completion of insert operation to the insert destination (2) to the insert data, issues an insert execution request to the insert destination (1) (step S1809), and, after the associated DBMS executor completes the insert operation, receives its result (step S1810).

In the example of Fig. 20, when the hash function values are 0, 9, 4 and 6, the insert destination (1) is the same as the insert destination (2), that is, the same storage 620a. When the hash function values are 1, 8 and 5, the inserts (1) and (2) are the same as the insert storage 620b. Thus the DBMS acceptance section issues an insert request to the respective DBMS executors 610a and 610b and receives their insert results.

When the hash function value is 7, the insert

5

10

20

25

5

10

15

20

25



5           Next, in the insert operation 2102 of such  
data 2131 as to provide a hash function value of 3  
during the rebalance operation, the data 2141 is stored  
in the storage 2140 and the data 2131 added with the  
data position information 2132 is stored in the storage  
0 2130.

When another user tries to perform search/update/delete operation over the data being moved during the rebalance operation, he has to wait until the move operation of the data is completed because the data is locked to prohibit access thereto. However once the lock is release due to the move completion of the data, his operation so far awaited can be resumed.

Further, even when the rebalance operation  
20 tries to move the data being searched/updated/deleted,  
the lock of the data causes prohibition of access to  
the data. Thus the move operation has to be awaited  
until the user release the lock. Once the lock is  
released, the rebalance operation so far awaited can be  
25 resumed. Thus it can be prevented that the value of  
the data be unduly updated or the data be deleted.

Further, since the data moved through the rebalance operation remains in the existing areas even

In this manner, in the database system of the

5 second embodiment, the data moved by the rebalance operation is added. Therefore, when the update/delete operation is executed over the data indicated by the position information at the time of updating/deleting the data added with the position information, the  
10 update/delete operation can be realized even over the data moved by the rebalance operation.

Further, when data to be rebalanced is stored in the same format as data being rebalanced, generation of excessive rebalance operation can be prevented.

15           As a result, the system can accept search,  
update, delete and insert processing requests to the  
table being rebalanced during the rebalance operation  
and can execute them concurrently.

In the database management system of the arrangement allowing realization of both of the aforementioned first and second embodiments, when the function operation of either one of the first and second embodiments is selected on the basis of the presence or absence of a column type or index included in the table, the rebalance operation and the search, update, delete and insert operations optimum for the attribute of the table can be executed.

Similarly, it is also possible to reduce the

number of storages for data storage.

In the database management system of Fig. 23, such storage areas of a stock table as shown in Fig. 24 and the storage 620c is disconnected. Thus when the storage C is defined in a disconnection area information 636 of the storage 630 and a rebalance processing request is generated, the operation of a rebalance operation executor 1101 is started.

The details of the rebalance operation of the  
10 rebalance operation executor 1101 are nearly the same  
as when the aforementioned explanation based on Fig.  
11, but data to be moved corresponds to all the data  
stored in the disconnection area. Further, the number  
of area divisions is reduced from three to two so that,  
15 on the basis of the hash function value storage  
position correspondence table 635, data indicative of  
hash function values of 2, 7 and 3 stored in the  
storage 620c and having a hash function value of 7 is  
moved to the storage 620a, and data indicative of hash  
20 function values of 2 and 3 is moved to the storage  
620b.

When the rebalance operation is completed, the data are stored in the storages 620a and 620b as shown in Fig. 25.

25           The operation when data search, update or  
delete request is issued during the rebalance operation  
is nearly the same as the aforementioned description.  
The operation when an insert request is issued during



the rebalance operation is nearly the same as the  
aforementioned description, except for the operation of  
the insert destination decider.

The operation of the insert destination  
5 decider find a hash function based on data to be  
inserted, refers to the hash function value storage  
position correspondence table 635, and selects an  
insert destination from one of the areas defined in the  
storage area information 632 by the hash function other  
10 than the area defined in the disconnection area  
information 636.

As has been explained in connection with  
Figs. 1 to 25, the database system and database  
management method in this example performs its data  
15 management operation according to a procedure which  
follows.

(a) In the database system for allocating the  
data of the table having a plurality of data items to a  
plurality of storages and storing the table data in  
20 storages determined by a specified division rule  
(method) such as, e.g., hash function, when a storage  
is required to be added for storage of the table data,  
the system determines data to be moved to the  
additional storage according to the hash function with  
25 use of information on the existing and additional  
storages, performs the data rebalance operation to move  
all the determined data to the additional storage.  
Further, when a search, update or delete request to the

09587839-4464  
10944-63828660

5

10

15

20

25

(c) When a storage is required to be added for storage of the table data, the system first copies data to be determined as moved according to the division rule using the hash function or the like from the existing storage to the additional storage, and adds position information of the copied data on the additional storage to the data of the storage as the copy source. And when completing the copying operation of all the data to be moved in this way, the system executes the data rebalance operation to delete all the data of the copy source. Further, when a search request to the table data is issued during the execution of the data rebalance operation, the system performs search operation over the data stored in the existing storages. When an update or delete request is issued, the system performs update or delete operation over the data stored in the existing storages. When the data as the update or delete object is added with position information indicative of a copy destination to the additional storage, the system performs the same update or delete operation over the data of the copy destination. Further, when an insert request to the table data is issued, the system stores the data in the storage determined according to the division rule

specified only for the existing storages. Further, in a division method which is applied to the storages including the additional storage, when the data is to be moved to the additional storage, the system stores  
5 the data even in the additional storage and adds storage position information thereof to the additional storage to the data stored in the existing storages.

Any one of the above techniques (a) to (c) is selected and executed depending on a table to be  
10 processed.

In this way, the database system and database management method in this example, at the time of performing rebalance operation over the data stored in the existing storages due to addition of a storage to  
15 be rebalanced in table data of a relational database management system of storing as divided into a plurality of storages, can execute database services such as search, update, delete and insert concurrently.

The present invention is not limited to the  
20 embodiments explained with reference to Figs. 1 to 25 but may be modified in various ways so long as it does not depart from its gist or subject matter. For example, although the database system and database management system have used three storages in the  
25 present embodiment, the number of such storages may be two or four or more. Even with regard to the data division rule (method) to the respective storages, the present invention is not limited only to the use of the

0987839-11604

In the present embodiment, the DBMS 300 has been configured with computers as shown in Fig. 2 as its hardware configuration. However, like the user terminals 330a to 330c, the DBMS may be configured with computers provided with input devices such as keyboards and display devices such as CRT (cathode ray tube).

10 Though explanation has been made in connection with the  
case where an optical disk is used as the recording  
medium in the present embodiment, a flexible disk (FD)  
may be used as the recording medium. Even with respect  
to program installation, the program can be downloaded  
15 from a communication device via a network and then  
installed.